



Curso: ¿Cómo se maneja SPSS como un profesional? Introducción al trabajo con sintaxis

José Manuel Torrado
Profesor Ayudante Doctor
Departamento de Sociología, Universidad de Granada
josetr@ugr.es

LECTURA 3: ANÁLISIS DE DATOS CON SPSS

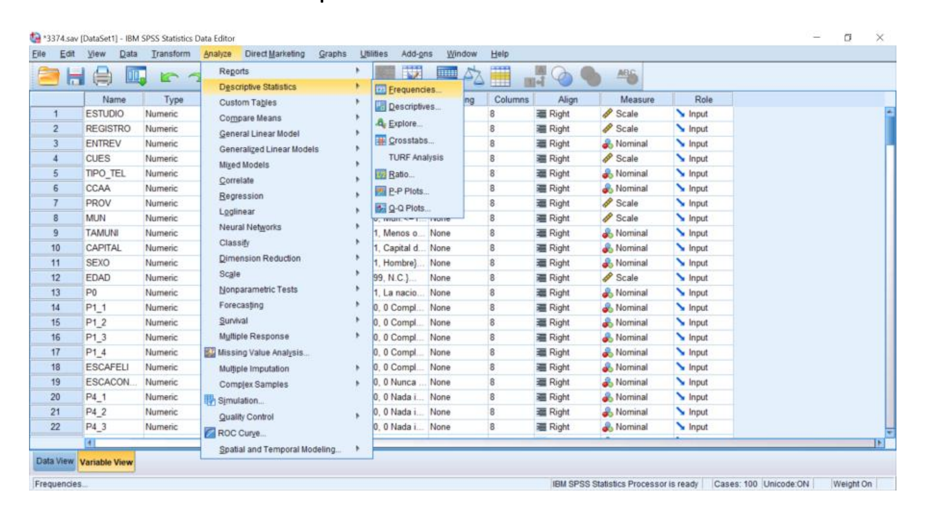
Hasta aquí hemos visto: En la lectura 1, el proceso de investigación social cuantitativa y dentro del mismo el proceso de análisis de datos y su fases, estas eran, recordemos: el análisis descriptivo univariable, el análisis descriptivo bivariable y el análisis inferencial (recomiendo vayan al apartado "análisis de datos" para recordar mejor en qué consiste cada etapa, ya que en este documento se dan por supuestas). En la lectura 2, os explicado cómo trabajar con datos secundarios facilita mucho el proceso de investigación social cuantitativo, y también se muestra cómo podéis descargar fuentes del CIS. En esta lectura 3 vamos a ver, paso a paso, cómo realizar el proceso de análisis de datos con SPSS.

Trabajando con la interfaz de usuario

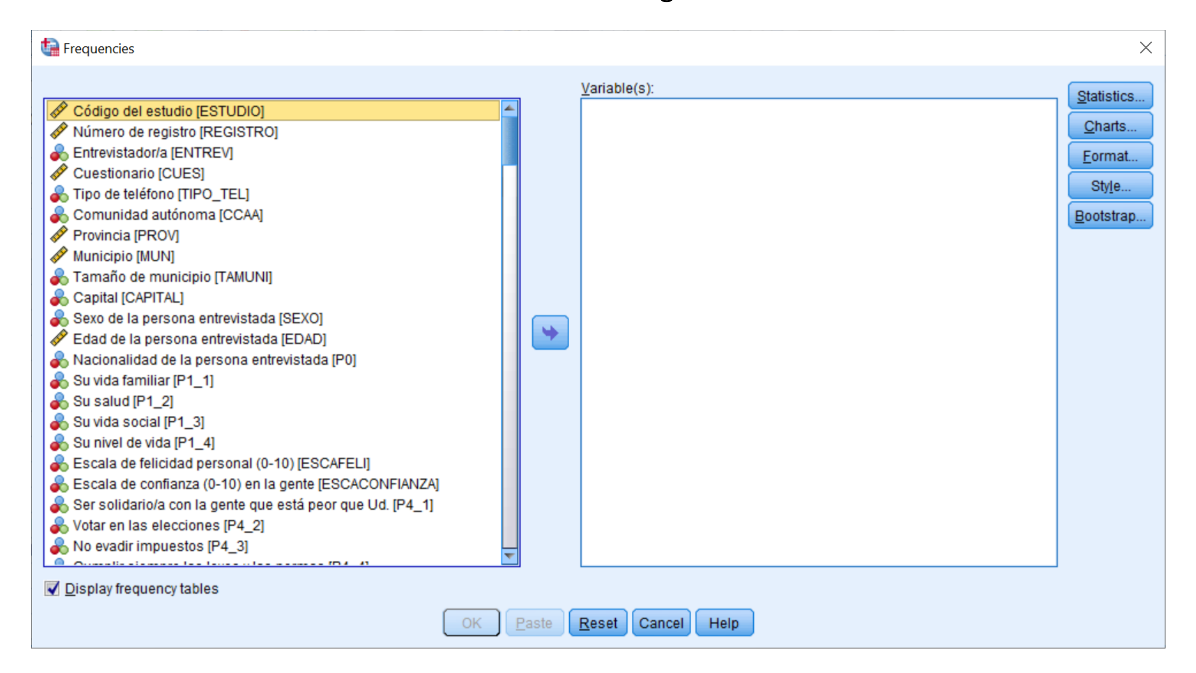
La forma más básica de trabajar con SPSS – y en la que se os forma en esta Facultad – es trabajando con la interfaz de usuario del programa, es decir, haciéndole peticiones al programa clicando en las distintas opciones que os da. Si bien esta es la forma más intuitiva, no es la más profesional ni la más recomendable. No obstante, pensando especialmente en los más novatos, aquí explicaré como se hace, para después ir un paso más allá y empezar a trabajar con sintaxis.

Análisis de tablas de frecuencias

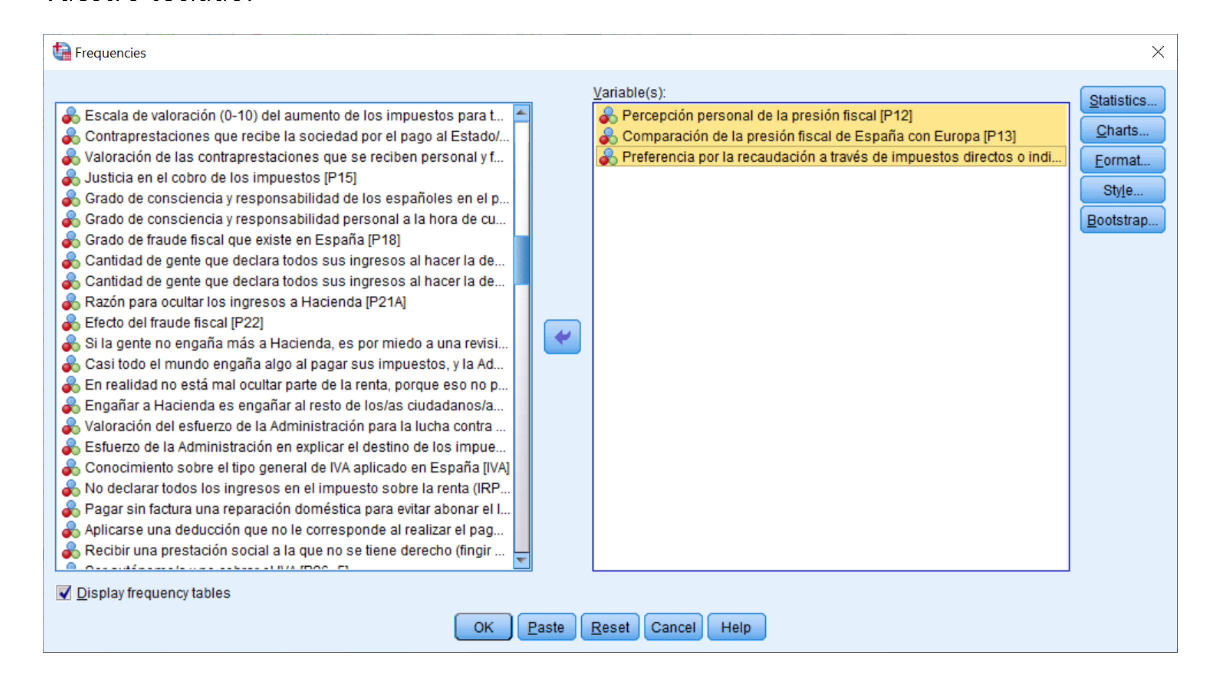
El primer paso en el proceso de análisis de datos es obtener las tablas de frecuencias de las principales variables que queremos analizar. Para ello, tenéis que clicar en "Análisis" → "Estadísticos descriptivos" → "Frecuencias"



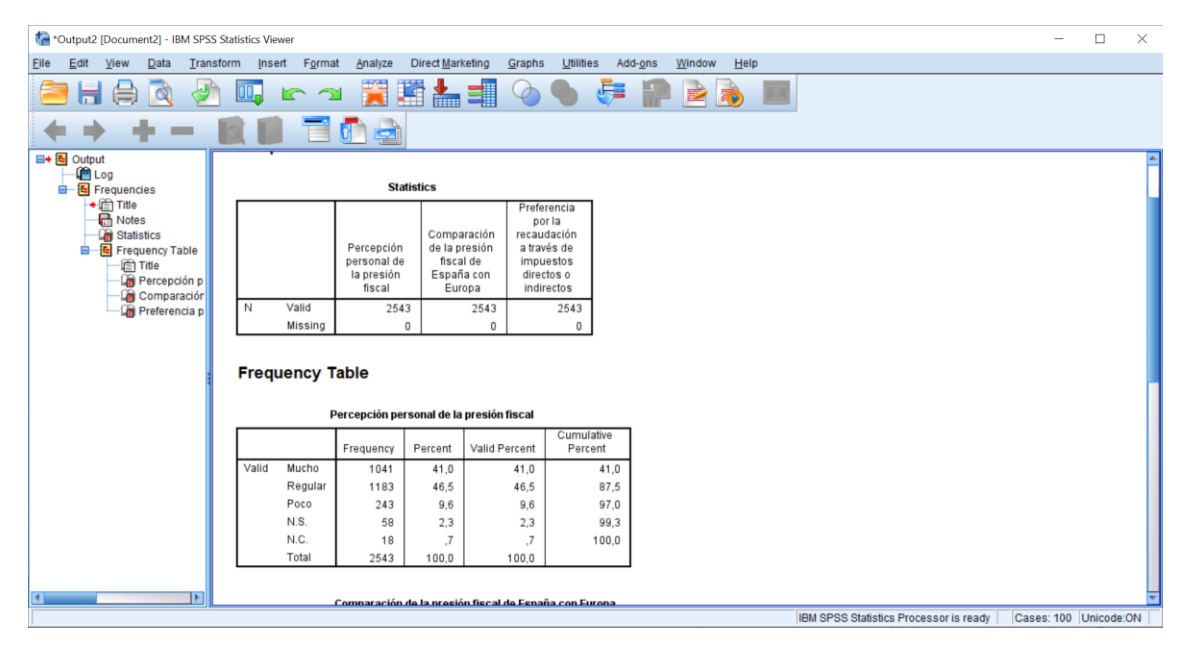
Una vez clicáis en “Frecuencias” se os abrirá la siguiente ventana:



Esta ventana contiene, a la derecha, todas las variables que tenga la base de datos, y a la izquierda, aquellas de las que queréis obtener su tabla de frecuencias. Para arrastrar variables a la izquierda debéis seleccionarla y clicar en la flecha . Podéis seleccionar varias variables de la izquierda si mientras clicáis mantenéis pulsada la tecla “Ctrl” de vuestro teclado.



Una vez seleccionadas las variables, le dais al botón para que el programa os de las tablas de frecuencias de las variables seleccionadas. Estas se abrirán en una ventana de “resultados” aparte de la ventana de trabajo, tal y como podéis ver en la siguiente captura.



Statistics

		Percepción personal de la presión fiscal	Comparación de la presión fiscal de España con Europa	Preferencia por la recaudación a través de impuestos directos o indirectos
N	Valid	2543	2543	2543
	Missing	0	0	0

Frequency Table

Percepción personal de la presión fiscal

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Mucho	1041	41,0	41,0	41,0
	Regular	1183	46,5	46,5	87,5
	Poco	243	9,6	9,6	97,0
	N.S.	58	2,3	2,3	99,3
	N.C.	18	,7	,7	100,0
	Total	2543	100,0	100,0	

Comparación de la presión fiscal de España con Europa

Lo primero que aparece (que no se ve en la captura), es el código de programación o sintaxis que el programa ha ejecutado para obtener lo que le habéis pedido (en este caso tabla de frecuencias), a continuación os da una tabla "Statics" en la que os dicen cuántos casos "válidos" tiene las variables que estáis analizando y cuántos "perdidos", por último aparecen la "Tablas de frecuencia" de cada una de las variables. Como vimos en la lectura 1, la tabla de frecuencia lo que indica es la cantidad de personas que han contestado a cada una de las categorías de respuesta posible (campo "Frecuencia"), qué porcentaje representan (campo "Porcentaje"), qué porcentaje representan respecto a los perdidos (campo "porcentaje válido") y el porcentaje acumulado (que es la suma del porcentaje de una categoría más la anterior). Me conformo con que sepáis leer los dos primeros campos. Por ejemplo, la tabla que os pego abajo, lo que nos dice es que 1041 personas perciben que pagan muchos impuestos, y estos representan un 41% del total de los encuestados.

Percepción personal de la presión fiscal

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Mucho	1041	41,0	41,0	41,0
	Regular	1183	46,5	46,5	87,5
	Poco	243	9,6	9,6	97,0
	N.S.	58	2,3	2,3	99,3
	N.C.	18	,7	,7	100,0
	Total	2543	100,0	100,0	

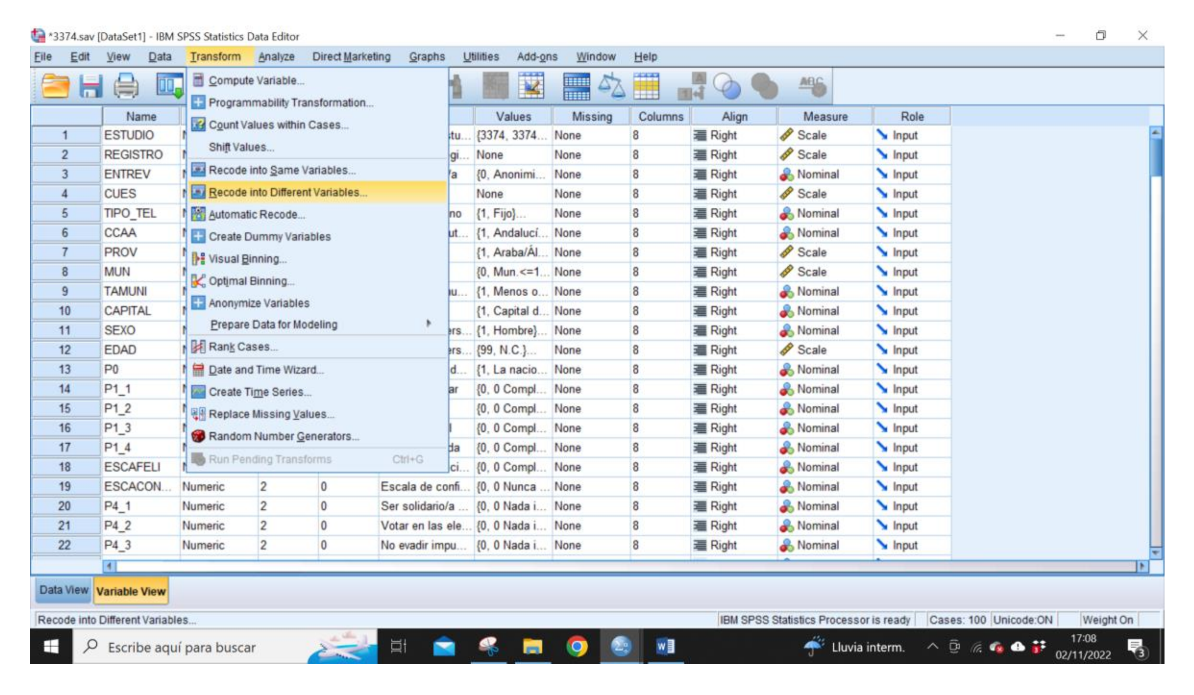




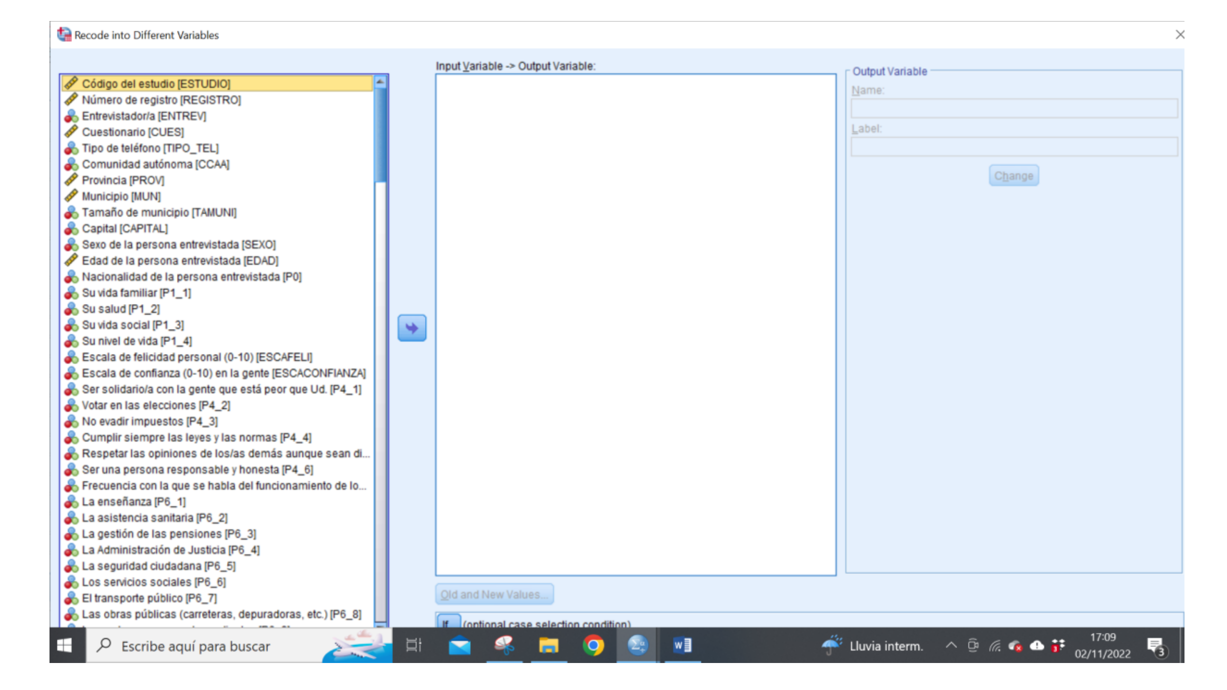
Recodificaciones en distintas variables

Una vez analizadas las variables que nos interesan, podemos tomar la opción, si lo creemos necesario, de recodificarlas. El programa ofrece dos opciones para esto: recodificar en la misma variable o recodificar en distintas variables. Yo soy partidario de hacerlo siempre en distintas variables, ya que de esta manera no perdemos la variable original (de la otra forma reescribiría la información sobre la misma variable, perdiendo para siempre la información de la variable original).

Para hacer esto, clicamos en "Transformar"→ "Recodificar en distintas variables"

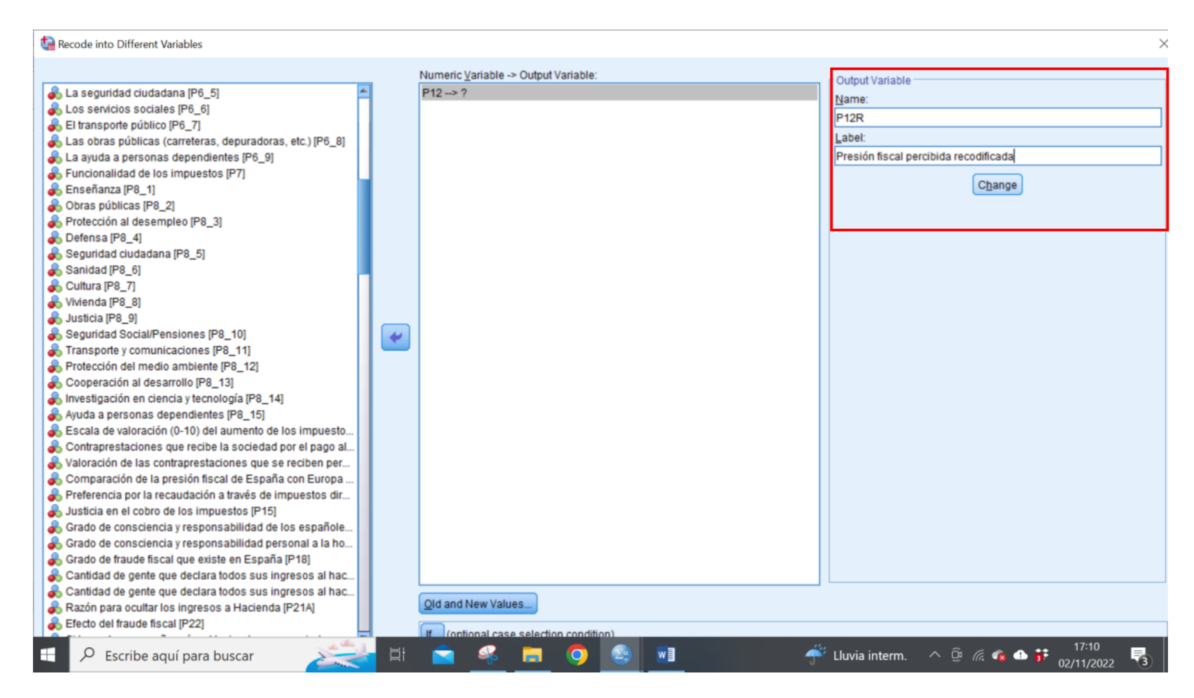


Una vez clicamos se nos abrirá la siguiente ventana:

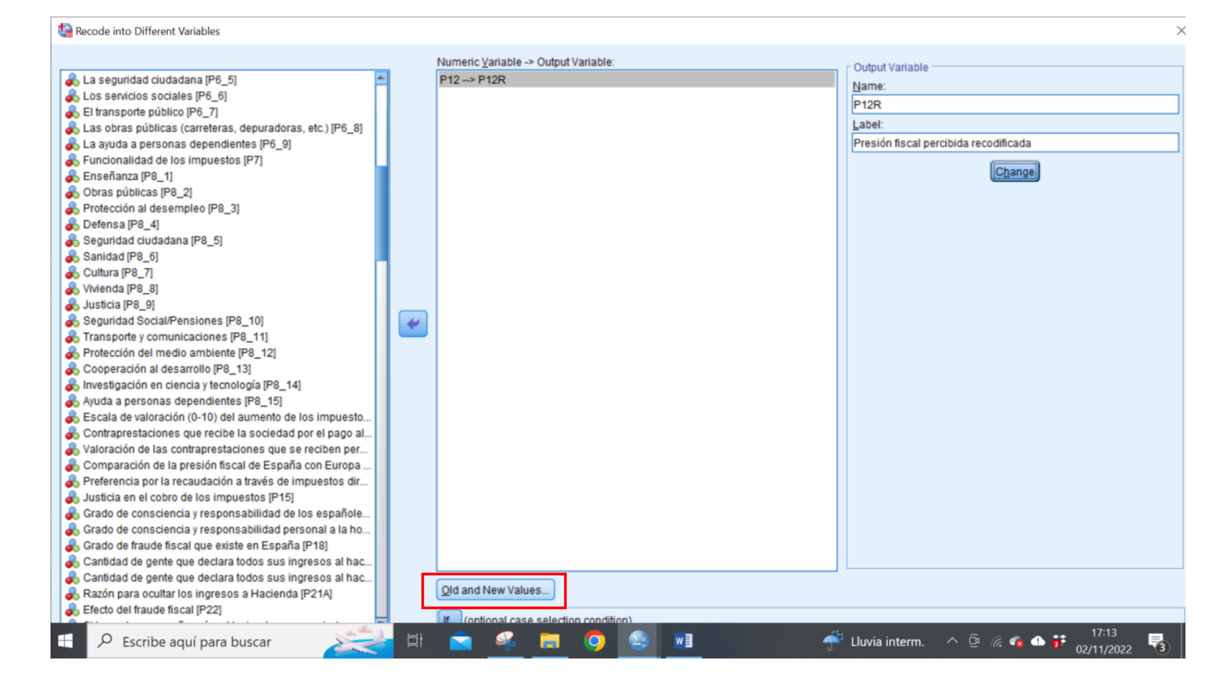




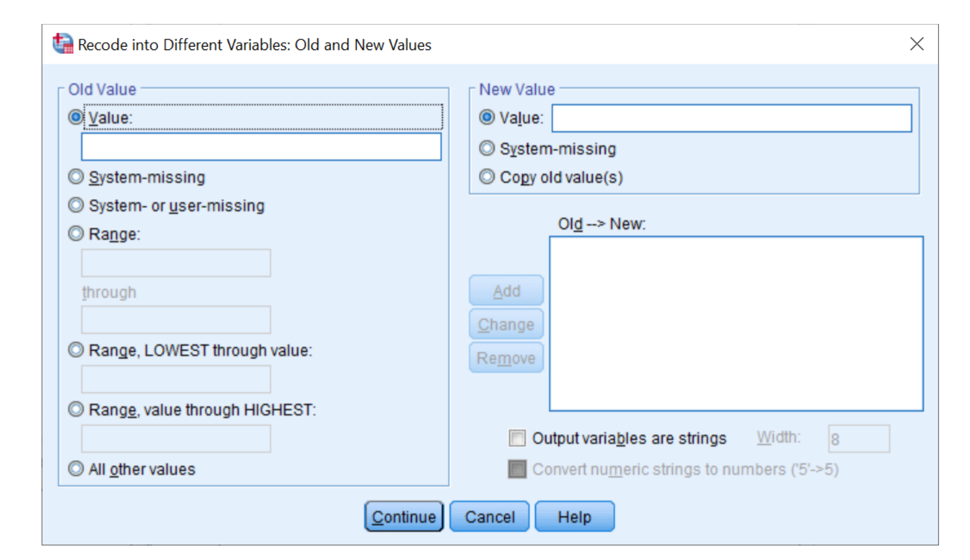
De nuevo, a la izquierda tenemos las variables de la base de datos con la que estemos trabajando, a la derecha, las que vamos a transformar. Voy a hacer la recodificación de la variable cuya tabla de frecuencias he mostrado antes, concretamente voy a unir los "NS" y "NC" en un misma categoría. Primero selecciono la variable, en el campo "Nueva variable" (recuadro señalado en rojo), pongo el nombre de la nueva variable y, escribo su descripción en el campo "etiqueta". Clicamos en "Cambiar" para que se ejecute el cambio.



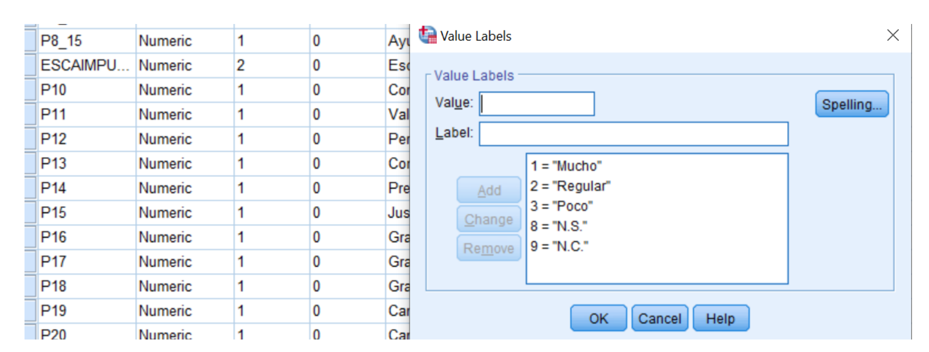
Hecho esto, clicamos en "Valores antiguos y nuevos" (recuadro señalado en rojo en la siguiente captura).



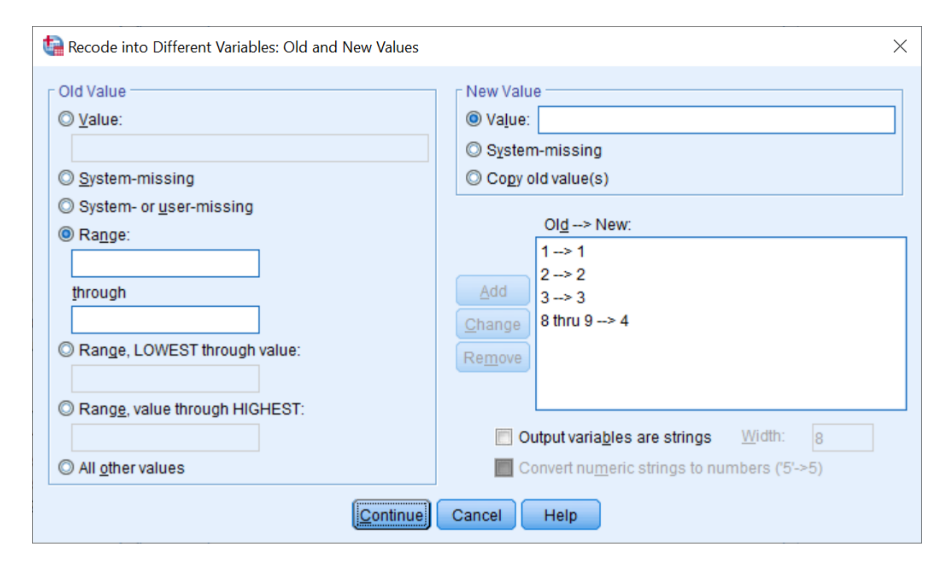
Una vez en “valores antiguos y nuevos” se nos abrirá la siguiente ventana:



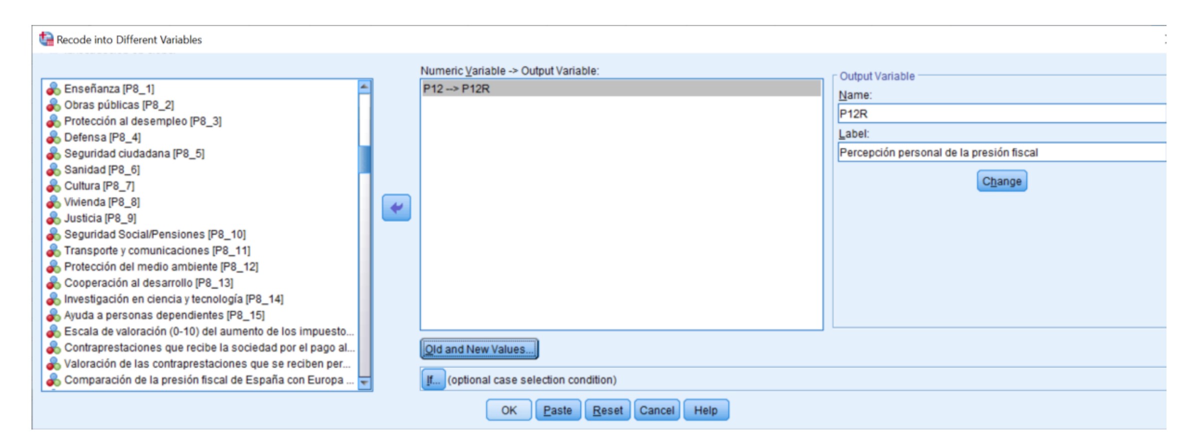
En la izquierda tenemos que escribir los valores antiguos que queremos recodificar, y en la derecha se escribe el valor nuevo que asumirá. Para hacer esto tenéis que ir a “vista de variables” para saber con certeza qué valor numérico tiene asignado cada categoría. En nuestro ejemplo son:



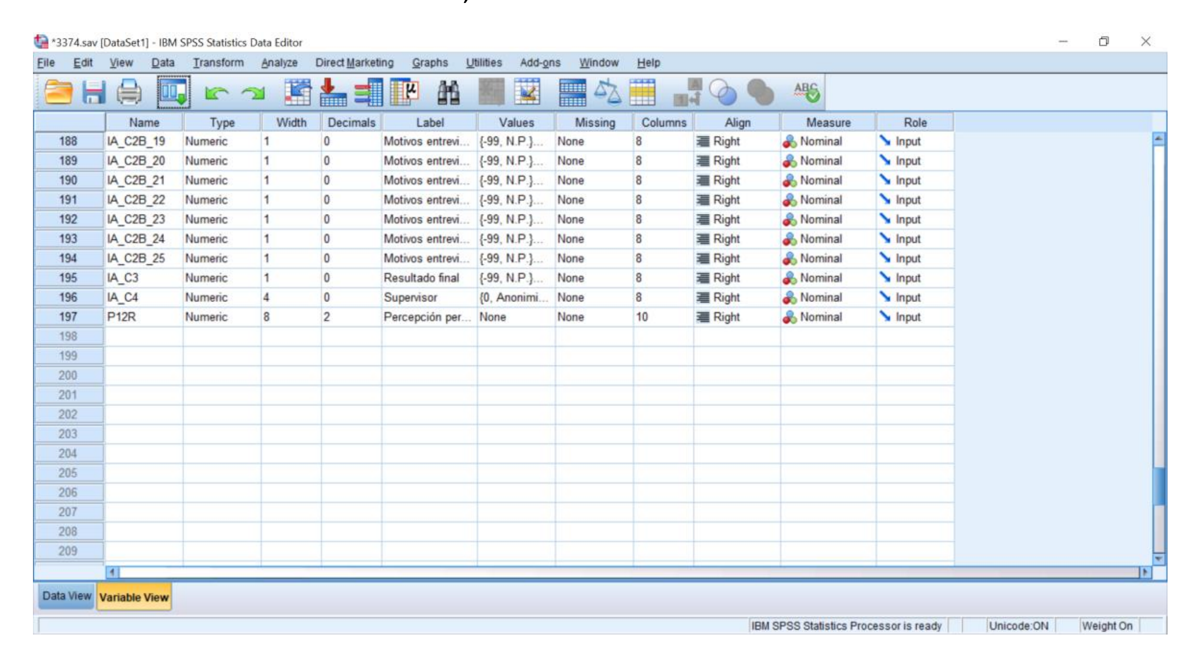
Con esta información ya podemos hacer los cambios oportunos:



Introducidos los cambios le damos a continuar, volveremos a la anterior ventana

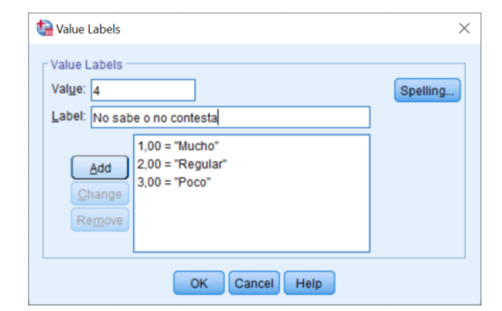


Le volvemos a dar a “OK” y ya tenemos creada la nueva variable (P12R). La podemos encontrar en “Vista de variables”, al final de la lista:



	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
188	IA_C2B_19	Numeric	1	0	Motivos entrevi...	{-99, N.P.}...	None	8	Right	Nominal	Input
189	IA_C2B_20	Numeric	1	0	Motivos entrevi...	{-99, N.P.}...	None	8	Right	Nominal	Input
190	IA_C2B_21	Numeric	1	0	Motivos entrevi...	{-99, N.P.}...	None	8	Right	Nominal	Input
191	IA_C2B_22	Numeric	1	0	Motivos entrevi...	{-99, N.P.}...	None	8	Right	Nominal	Input
192	IA_C2B_23	Numeric	1	0	Motivos entrevi...	{-99, N.P.}...	None	8	Right	Nominal	Input
193	IA_C2B_24	Numeric	1	0	Motivos entrevi...	{-99, N.P.}...	None	8	Right	Nominal	Input
194	IA_C2B_25	Numeric	1	0	Motivos entrevi...	{-99, N.P.}...	None	8	Right	Nominal	Input
195	IA_C3	Numeric	1	0	Resultado final	{-99, N.P.}...	None	8	Right	Nominal	Input
196	IA_C4	Numeric	4	0	Supervisor	{0, Anonimi...	None	8	Right	Nominal	Input
197	P12R	Numeric	8	2	Percepción per...	None	None	10	Right	Nominal	Input
198											
199											
200											
201											
202											
203											
204											
205											
206											
207											
208											
209											

Sin embargo, esta variable nueva no tiene etiquetas en los valores. Para etiquetarlos, en la misma “vista de datos”, clicamos en el campo “valores” de la nueva variable. Que aparecerá vacío y escribimos qué significa cada valor numérico asignado. Una vez añadidos todos, ya tenemos lista la nueva variable.

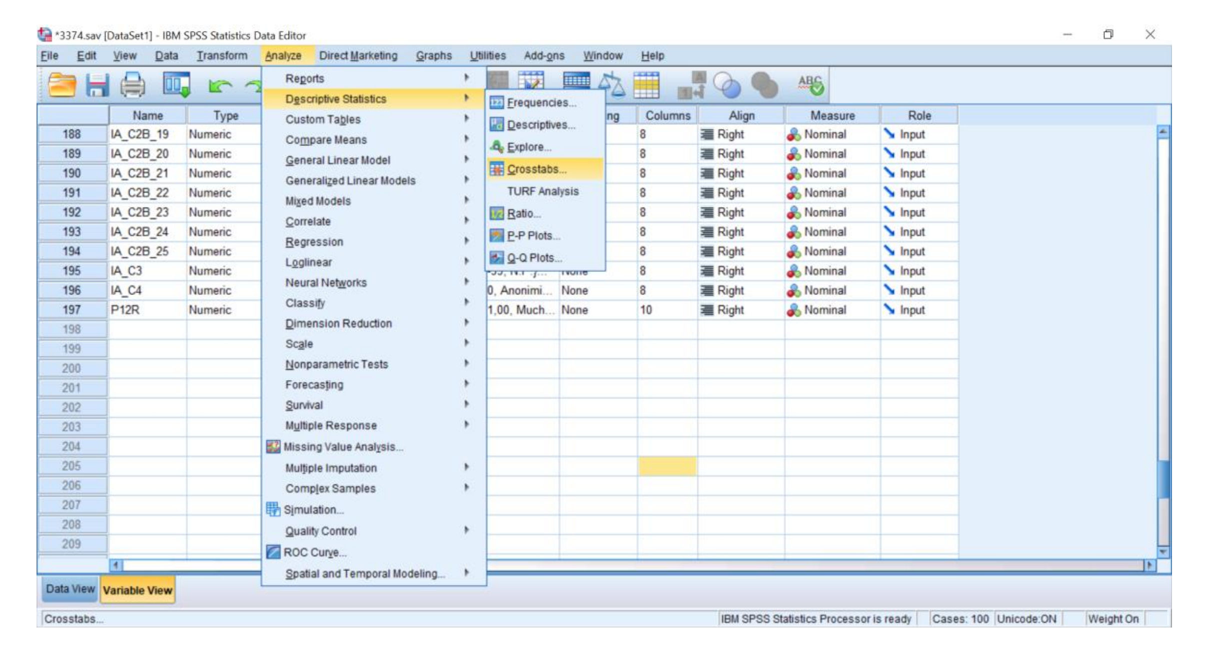




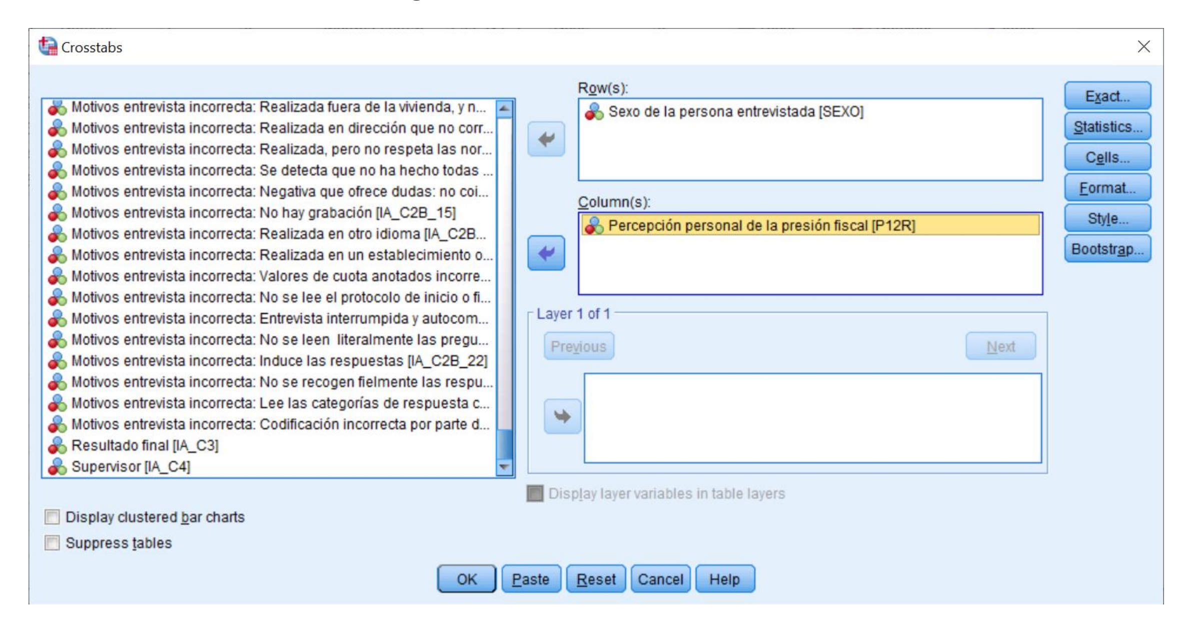


Tablas de contingencia

Una vez realizado el análisis de frecuencias y recodificadas las variables que consideremos oportunas, toca pasar al análisis bivariable, para ver cómo las opiniones varían entre los distintos grupos sociales. Para hacer una tabla de contingencia en SPSS debemos clicar en "Analizar" → "Estadísticos descriptivos" → "Tablas de contingencia"

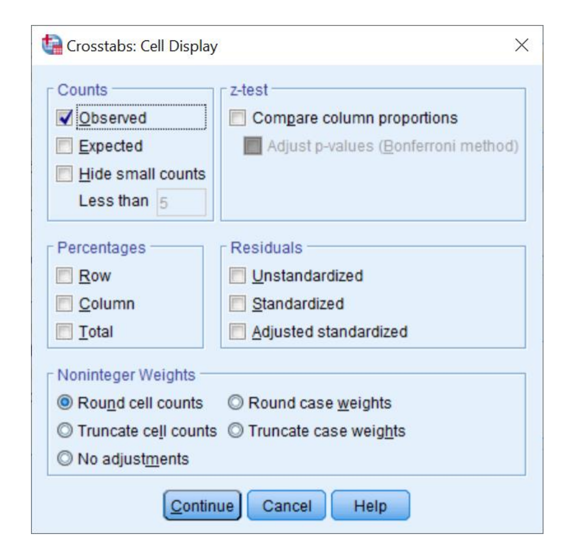


Una vez ahí se nos abrirá la siguiente ventana:

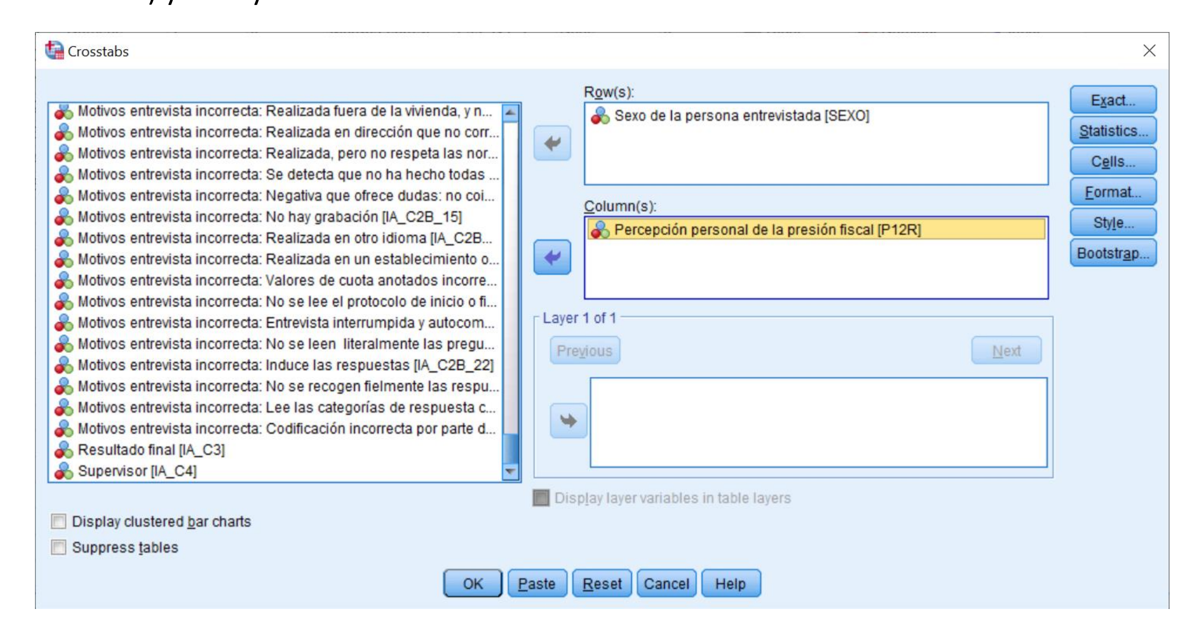


En las columnas metemos la variable que queramos explicar, en este ejemplo he metido la variable que hemos creado "percepción personal de la presión fiscal" y en las filas la variable que creemos que la explica, en este caso he metido la variable "sexo". Antes de darle a OK, vamos a "Casillas" (en mi caso "cells" porque tengo el programa en inglés). Se os abrirá la siguiente ventana:





Como veis por defecto solo viene señalada la opción "Observada" ("observed" en inglés), en este caso, también me interesa que me de los porcentajes de file, para conocer la distinta propensión de hombres y mujeres a percibir más o menos presión fiscal. Para ello en "porcentajes" ("percentages" en mi caso), clico en "filas" ("rows" en mi caso") y le doy a OK. Me llevará a la ventana anterior:



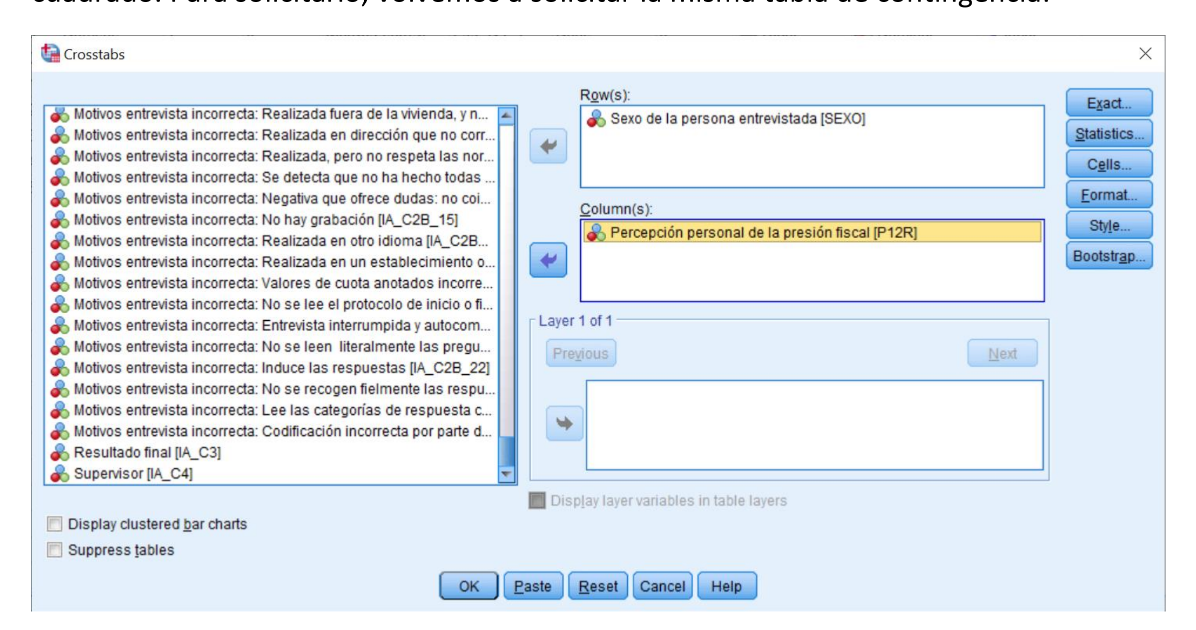
Le damos a OK, y en la ventana de resultados aparecerá la tabla que le hemos pedido:

Si analizamos la tabla vemos que no hay apenas diferencias entre hombres y mujeres, ya que los porcentajes son muy similares, por lo que podemos concluir que el sexo no influye en la percepción de la presión fiscal.

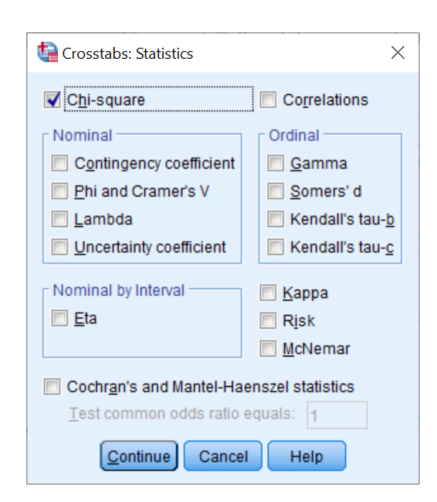
Análisis inferencial: contraste de hipótesis con Chi-Cuadrado

Por último, una vez hemos analizado todos los cruces de variables que consideramos pertinentes, cabría preguntarse por si las diferencias encontradas en alguno de ellos es

estadísticamente significativa. Para contrastarla analizaremos el coeficiente Chi-cuadrado. Para solicitarlo, volvemos a solicitar la misma tabla de contingencia:



Clicamos en la opción "Estadísticos" ("Statics" en mi caso), accediendo a la siguiente ventana:



Clicamos en Chi-cuadrado y le damos a continuar. Nos devolverá a la ventana anterior, le damos a OK. Y, en la ventana de resultados, obtenemos la siguiente tabla:

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	10,673[a]	6	,099
Likelihood Ratio	10,712	6	,098
Linear-by-Linear Association	6,788	1	,009
N of Valid Cases	1409		

Fuente: CIS 2022, Encuesta Flash Senado




¿Cómo se interpreta esto? Es sencillo. Si la significación asintótica (Asymptotic Significance) del "Pearson Chi-Square" es menor o igual que 0,05, entonces diremos que la correlación entre las variables es significativa, y en nuestro caso, que el sexo explica la percepción de la presión fiscal. Si ese valor es mayor que 0,05 entonces diremos que no es significativa, y por tanto no podremos extrapolar la interpretaciones que hagamos de la tabla (por ejemplo diciendo que las mujeres tienen más probabilidad a percibir una mayor presión fiscal) al conjunto de la población, como es nuestro caso, al ser ese valor de 0,099, o sea, mayor que 0,05 y por tanto, no significativo.